



# Ελληνοποίηση Πλατφόρμας Talend Open Studio

Απολογισμός και προοπτικές

Ομάδα Έργου:

Φαρδέλας Κωνσταντίνος

Κορύλλος Κωνσταντίνος

# Ποιοί είμαστε?

Λίγη προπαγάνδα....!



telecompare™  
σύμβουλοι τηλεπικοινωνιών



Τι είναι το Talend Open  
Studio;



# A versatile data integration solution

ETL (Extract – Transform - Load)

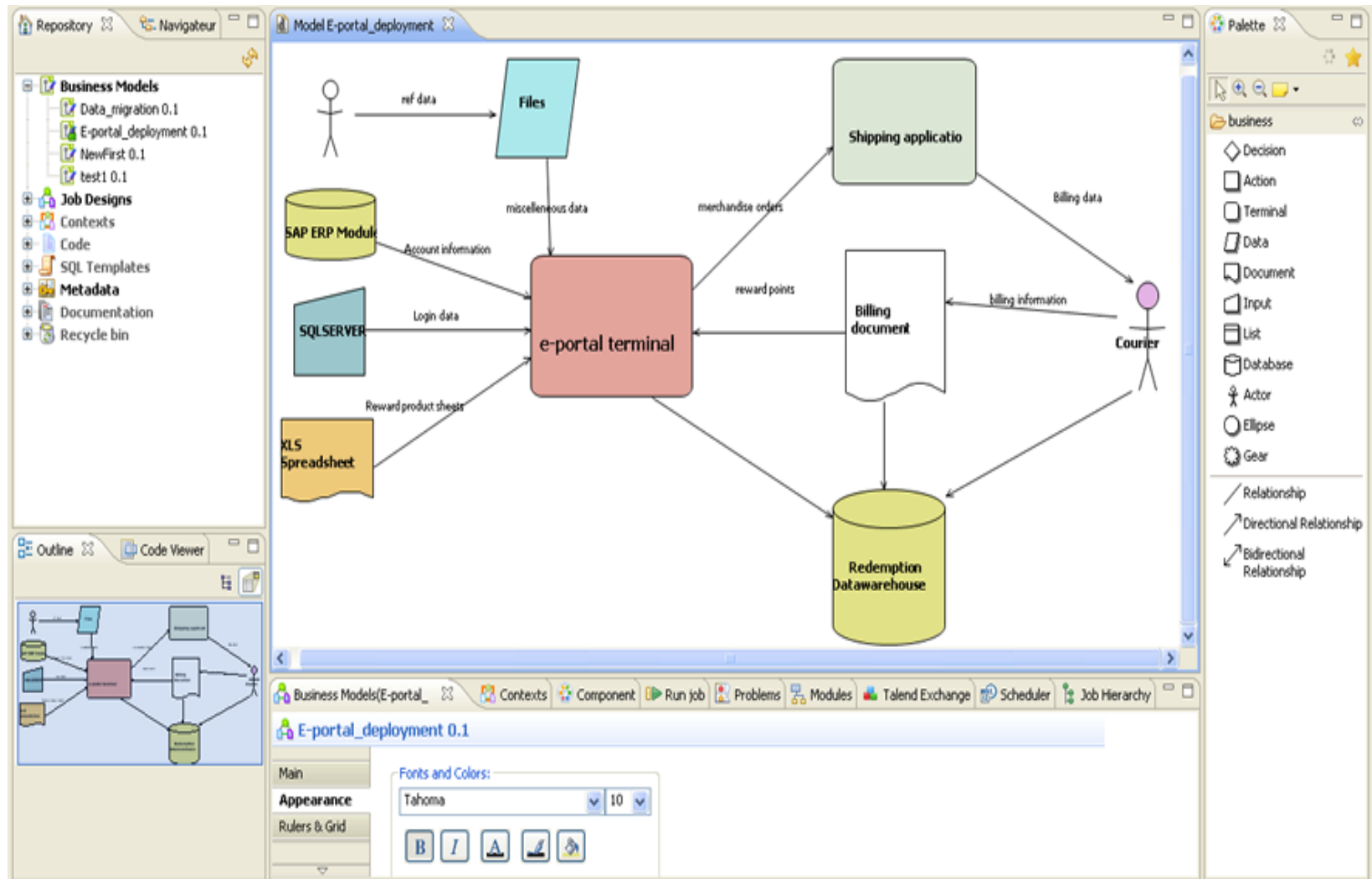
Συγχώνευση Δεδομένων (Data migration)

Μεταφορά Δεδομένων μεταξύ διαφορετικών συστημάτων

Συγχρονισμός Βάσεων Δεδομένων

Μετατροπή Δεδομένων και  
Μεταφόρτωση

# Business oriented process modelling



# Γραφικό Περιβάλλον...

The screenshot displays the Talend Studio graphical user interface (GUI) for designing a data integration job. The main workspace shows a job design for 'Job tellas2gcf 0.1' with various components and connections. The components include file inputs/outputs, tMaps, tUnites, tFilters, and tJavaRow. The connections are labeled with row names and conditions. The left sidebar shows the 'Repository' tree with 'Business Models' and 'Job Designs' categories. The right sidebar shows the 'Palette' with various components categorized by function. The bottom of the screen shows the 'Designer' tab with the 'tFileList\_1' component selected, displaying its 'Basic settings' and 'Advanced settings'.

**Repository**

- Business Models
- Job Designs
  - Cosmote
  - eletson
  - HOL
  - palo
  - tellas\_wind
    - europisti\_tellas\_wind 0.1
    - europisti\_tellas\_wind\_final 0.1
    - tellas\_wind\_temp\_process 0.1
    - tellas2gcf 0.1
    - wind\_ebill\_to\_mob\_gcf 0.1
    - wind\_ebill\_to\_mob\_gif 0.1
    - wind\_mob\_gcf\_to\_csv 0.1
  - Vodafone
    - altec2gcf 0.1
    - dslams\_to\_gis 0.1
    - genmills\_pbx 0.1
    - image\_metadata\_xls2xml 0.1
    - loadnastros 0.1

**Job tellas2gcf 0.1**

Components and Connections:

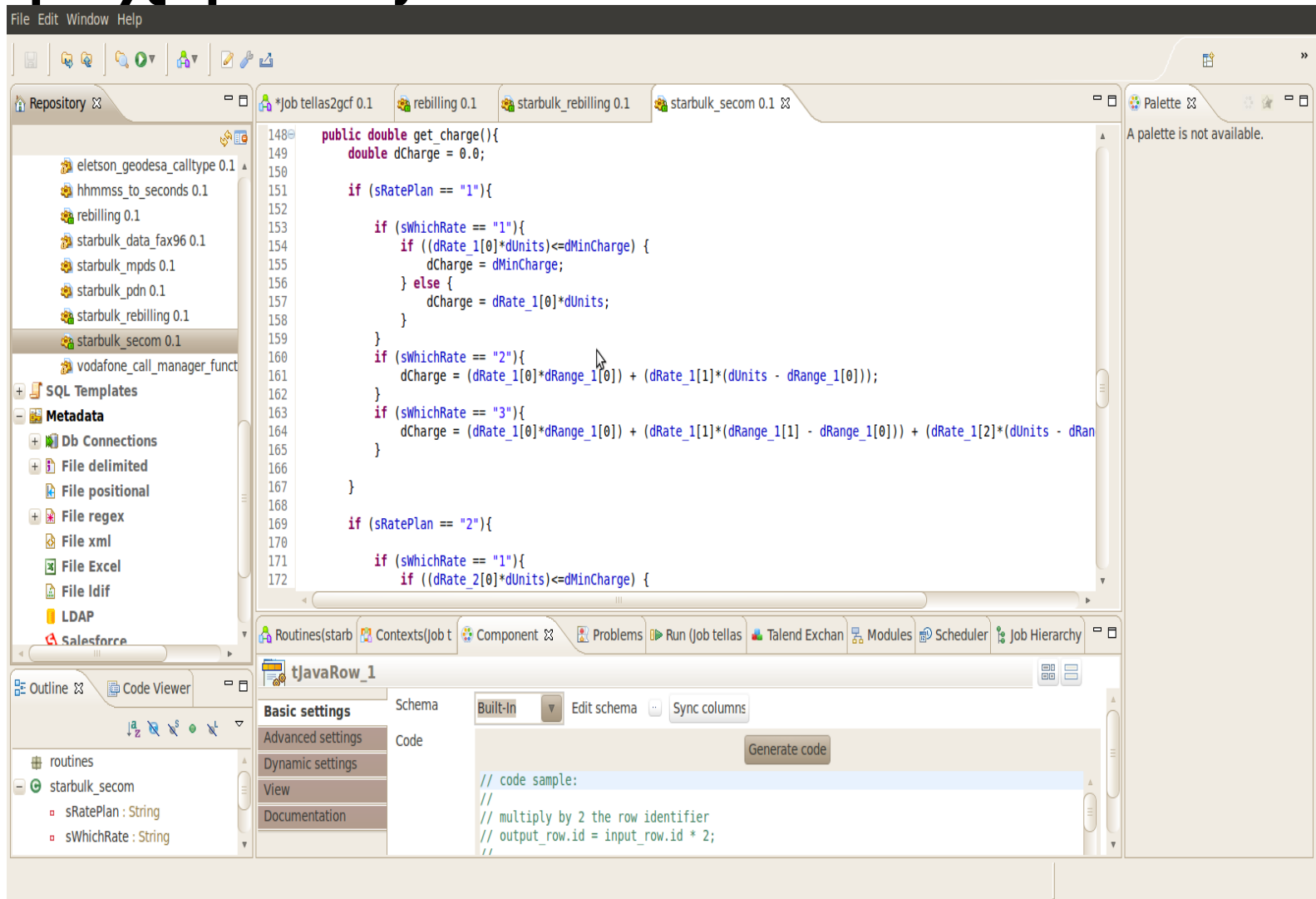
- tFileInputDelimited\_2 (Main) → tMap\_2 (Main) → tFileOutputDelimited\_3 (Main)
- tFileList\_1 (Main) → tMap\_2 (Main) → tFileOutputDelimited\_3 (Main)
- tFileInputDelimited\_1 (Main) → tMap\_1 (Main) → tFileInputDelimited\_4 (Main)
- tFileInputDelimited\_3 (Main) → tMap\_3 (Main) → tFileOutputDelimited\_4 (Main)
- tFileInputDelimited\_4 (Main) → tFilterRow\_2 (Main) → tFileOutputDelimited\_6 (Main)
- tFileInputDelimited\_4 (Main) → tFilterRow\_4 (Main) → tJavaRow\_1 (Main)
- tFileInputDelimited\_4 (Main) → tFilterRow\_11 (Main) → tJavaRow\_1 (Main)
- tFileInputDelimited\_4 (Main) → tFilterRow\_10 (Main) → tJavaRow\_1 (Main)
- tFileInputDelimited\_4 (Main) → tFilterRow\_7 (Main) → tJavaRow\_1 (Main)
- tFileInputDelimited\_4 (Main) → tFilterRow\_1 (Main) → tJavaRow\_1 (Main)
- tFileInputDelimited\_4 (Main) → tFilterRow\_11 (Main) → tJavaRow\_1 (Main)
- tFileInputDelimited\_4 (Main) → tFilterRow\_10 (Main) → tJavaRow\_1 (Main)
- tFileInputDelimited\_4 (Main) → tFilterRow\_7 (Main) → tJavaRow\_1 (Main)
- tFileInputDelimited\_4 (Main) → tFilterRow\_1 (Main) → tJavaRow\_1 (Main)

**tFileList\_1**

Basic settings

- FileList Type: Files
- Includes subdirectories: ☐
- Case Sensitive: Yes
- Generate Error if no file found: ☒
- Use Glob Expressions as Filemask (Unchecked means Perl5 Regex Expressions): ☒
- Files: Filemask: \*.csv

...μέ όσο κώδικα τραβάει η  
ψυχή σας...





# Metadata-driven design and execution

Metadata repository

Centralization

Consistency

Open relational model





# Advanced and versatile connectivity

**Files:** CSV, Excel, positional, XML, LDIF, unstructured

**Databases:** MySQL, MS SQL Server, DB2, Oracle, Teradata, Ingres, PostgreSQL, Sybase, MS Access, Informix, Firebird

**Applications:** SAP, SugarCRM, MS Dynamics, Salesforce.com

**Log & error:** info, warning, statistics, filtering

**Data quality:** deduplication, fuzzy logic matching, CRC

**Transformation:** filtering, mapping, reference, aggregation, lookup, XSLT

**Miscellaneous:** Web Services, FTP, HTTP, system, SSH



# Deployment and maintenance

## Real-time debugging

## Robust and scalable execution



# Εφαρμογή του Talend Open Studio στο Δημόσιο Τομέα



# Μια αξιόπιστη λύση

Μείωση Ολικού Κόστους Κτήσης (TCO)

Customization

Διαλειτουργικότητα

Ευκολία στη χρήση και στην εκμάθηση

Ισχυρή και δραστήρια κοινότητα (forum, wiki, tutorials, bugtracker etc)

Επιπλέον υπηρεσίες (Support, Training, Consulting)



# Στην πράξη


Μετάπτωση / Μεταφορά βάσεων δεδομένων  
(πχ. από Oracle / MSSQL σε PostgreSQL /  
MySQL)

Συγχρονισμό διαφορετικών συστημάτων  
μεταξύ τους – Interoperability

Εξαγωγή δεδομένων από πολλαπλά  
συστήματα και επεξεργασία τους με σκοπό  
την παραγωγή αναφορών (σε συνεργασία με  
BI πλατφόρμες)



# Παραδείγματα χρήσης από Δημόσιους Φορείς του εξωτερικού



## **Ινστιτούτο Προστασίας Ραδιοενέργειας και Πυρηνικής Ασφάλειας (IRSN)**

Συγχώνευση (integration) και προτυποποίηση εκμετάλλευσης δεδομένων ραδιοενέργειας από διαφορετικές πηγές

### **Η πόλη του Brantford**

Διαδικασίες συγχώνευσης (integration) και συγχρονισμού (synchronization) δεδομένων

### **Υπηρεσία Προσφύγων των Ηνωμένων Εθνών (UNHCR)**

Συγχώνευση δεδομένων και παραγωγή αναφορών σχετικά με την κατάσταση των προσφυγικών καταυλισμών



# Προοπτικές χρήσης στον Ελληνικό Δημόσιο Τομέα

Σχέδιο Καλλικράτης?!





# Περίληψη Έργου – Παραδοτέα



# Μετάφραση του user interface του λογισμικού

89% των λεκτικών του Talend Open Studio έχουν μεταφραστεί

Η 3η γλώσσα με πλήρη μετάφραση παγκοσμίως

22% - 46% των λεκτικών συγγενικών λογισμικών έχουν μεταφραστεί

23 νέοι όροι προς υποβολή στο Γλωσσάρι του ΕΛ/ΛΑΚ

Μετάφραση της τεκμηρίωσης

User guide (240 σελ.)

- <http://www.talendforge.org/forum/viewtopic.php?pid=41743#p41743>
- <http://www.talendforge.org/wiki/doku.php?id=start>

Components Reference Guide (861 σελ.)

- Δεν εντάχθηκε στο έργο και δε μεταφράστηκε λόγω περιορισμένου χρόνου



# Μέθοδος Εργασίας



# Μετάφραση του user interface του λογισμικού

<http://www.talendforge.org/babili/>

## Μετάφραση του User Guide

Απευθείας επεξεργασία των πηγαίων αρχείων (.fm αρχεία) μέσω του λογισμικού Adobe Framemaker v.8

[http://www.ellak.gr/index.php?option=com\\_glossary&Itemid=75](http://www.ellak.gr/index.php?option=com_glossary&Itemid=75)



# Προβλήματα

Χρήση εμπορικού λογισμικού σε έργο ανοιχτού λογισμικού!

Ελλιπής λειτουργικότητα του εργαλείου μετάφρασης των λεκτικών του λογισμικού

Απουσία σήμανσης στα λεκτικά που δεν πρέπει να μεταφραστούν

Ανεπαρκής επαλήθευση της ορθότητας των μεταφράσεων πριν την έκδοση του λογισμικού



# Για το μέλλον

Δημοσίως διαθέσιμα πηγαία αρχεία  
για μελλοντική επεξεργασία

<http://www.talendforge.org/wiki/doku.php?id=start>

Καταγραφή εκκρεμών σφαλμάτων και  
προτάσεις βελτιώσεων



Ευχαριστούμε πολύ!